# S(C)ENTINEL - Monitoring Automated Vehicles with Olfactory Reliability Displays*

**Philipp Wintersberger**
CARISSMA
Technische Hochschule Ingolstadt
Ingolstadt, Germany
Johannes Kepler University
Linz, Austria
philipp.wintersberger@carissma.eu

**Dmitrijs Dmitrenko**
SCHI Lab, Creative Technology Research Group, School of Engineering and Informatics, University of Sussex
Brighton, United Kingdom
d.dmitrenko@sussex.ac.uk

**Clemens Schartmüller**
Technische Hochschule Ingolstadt
Ingolstadt, Germany
Johannes Kepler University
Linz, Austria
clemens.schartmueller@thi.de

**Anna-Katharina Frison**
Technische Hochschule Ingolstadt
Ingolstadt, Germany
Johannes Kepler University
Linz, Austria
anna-katharina.frison@thi.de

**Emanuela Maggioni**
SCHI Lab, Creative Technology Research Group, School of Engineering and Informatics, University of Sussex
Brighton, United Kingdom
e.maggioni@sussex.ac.uk

**Marianna Obrist**
SCHI Lab, Creative Technology Research Group, School of Engineering and Informatics, University of Sussex
Brighton, United Kingdom
m.obrist@sussex.ac.uk

**Andreas Riener**
Technische Hochschule Ingolstadt
Ingolstadt, Germany
Johannes Kepler University
Linz, Austria
andreas.riener@thi.de

## ABSTRACT

Overreliance in technology is safety-critical and it is assumed that this could have been a main cause of severe accidents with automated vehicles. To ease the complex task of permanently monitoring vehicle behavior in the driving environment, researchers have proposed to implement reliability/uncertainty displays. Such displays allow to estimate whether or not an upcoming intervention is likely. However, presenting uncertainty just adds more visual workload on drivers, who might also be engaged in secondary tasks. We suggest to use olfactory displays as a potential solution to communicate system uncertainty and conducted a user study (N=25) in a high-fidelity driving simulator. Results of the experiment (conditions: no reliability display, purely visual reliability display, and visual-olfactory reliability display) comping both objective (task performance) and subjective (technology acceptance model, trust scales, semi-structured interviews) measures suggest that olfactory notifications could become a valuable extension for calibrating trust in automated vehicles.

*Both first and second author contributed equally

*IUI '19, March 17–20, 2019, Marina del Ray, CA, USA*
2019. ACM ISBN 978-1-4503-6272-6/19/03…$15.00
https://doi.org/10.1145/3301275.3302332

## CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**; *Interaction techniques*; Interaction paradigms.

## KEYWORDS

Automated Driving, Olfactory, Reliability, SAE J3016, Human Factors, Trust

**ACM Reference Format:**
Philipp Wintersberger, Dmitrijs Dmitrenko, Clemens Schartmüller, Anna-Katharina Frison, Emanuela Maggioni, Marianna Obrist, and Andreas Riener. 2019. S(C)ENTINEL - Monitoring Automated Vehicles with Olfactory Reliability Displays. In *Proceedings of 24th International Conference on Intelligent User Interfaces (IUI '19).* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3301275.3302332

## 1 INTRODUCTION

Trust in automation is an important topic for a safe use of automated driving systems (ADSs) [71]. According to the classification of autonomy levels as proposed by SAE [12], ADSs currently available on the market mainly operate on level 2. Here, the driver is fully responsible for monitoring the vehicle's actions and thus the overall safety. However, recent events (such as the fatal crash of a Tesla driver in May 2016, but also less critical situations) indicate that many drivers utilizing such systems tend to overtrust them, and do not properly monitor ADSs even in scenarios they were not designed for [67]. This is especially dangerous when systems seem to work flawlessly for a long time and in varying situations [21]. Since monitoring is a challenging task, even for *"highly motivated human beings"* (irony of automation [4]), researchers have proposed to use so-called "reliability/uncertainty displays", that have shown to provide benefits in both level 2 [6] and level 3/4 automated driving (AD) [25]. Such displays are able to reduce the chance of mode awareness failures while increasing situation awareness as well as system transparency, and thereby ultimately lead to better calibrated trust [49]. They present the actual system reliability (or uncertainty, what is the inversion of reliability, but still follows the same concept – which kind of information works better is still an ongoing research [49]) to the user to adjust his/her monitoring behavior. However, especially when drivers are visually engaged in secondary tasks, such displays can merely act as "proxy" for the system state – instead of monitoring the vehicle and the environment itself, the driver has to frequently inspect the display, what still demands his/her visual attention.

Since future intelligent and multimodal user interfaces should adapt to different types of users [51] while using the full range of human interaction and communication capabilities [63], we claim that there is a need to evaluate other modalities for communicating reliability information. A potential modality in this regard could be the sense of smell, that, in contrast to other typical approaches (such as haptics [39] or auditory cues [38]), is still widely unused, but provides some unique advantages: The sense of smell is a very powerful interaction medium [36] enabling humans to extract meaningful information [60]. For example, it has been shown that odors trigger automatic and implicit retrieval of mental representations of information related to the object the scent is coming from [11], and enable automatic access to terms semantically related to odors [30]. Moreover, scents can be very efficient in activating the central neural system [3, 37, 65], which is essential to keep the driver alert and more attentive on the road [56]. Scents can also act as an arousing (e.g., when the driver is tired or inattentive [23, 73]) or as a calming (e.g., when the driver is stressed [31, 46]) stimulus. In future automated vehicles (AVs), classical perception channels (i.e., visual and auditory) will often be occupied by secondary tasks (such as watching a video, what demands both visual and auditory attention), while olfactory notifications have proven to be a valid way to gain user's attention [7]. Consequently, to the best of our knowledge, our study (see Figure 1) is the first experiment including olfactory notifications for trust calibration in AD.

## 2 RELATED WORK

Trust in automation can be defined as *"the attitude that an agent will achieve an individual's goals in a situation characterized by uncertainty and vulnerability"* [41], and is a complex construct built by analytic, analogical, and affective processes before (dispositional trust), during (situational trust), and after (learned trust) direct system interaction [29]. To foster safe use of automated systems (and thereby prevent both disuse and misuse [52]), users should adjust their subjective trust levels to fit *"an objective measure of trustworthiness"* (*"calibration of trust"* [47]). Reliability/uncertainty displays should assist in the process of trust calibration (especially to account for overtrust) by providing decision aids that allow users to estimate an automated system's performance in a given situation [22].

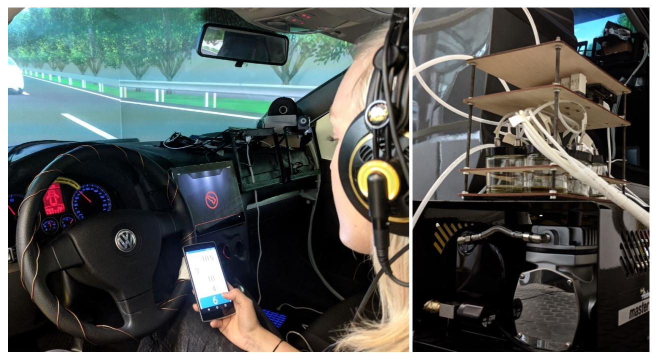

**Figure 1: Study setup: Participants had to frequently intervene by actuating the brake pedal in case the automated longitudinal system fails (low reliability indicated on the central in-vehicle display), while performing a detection-response task on a smartphone (left). To hide the activation sound of the olfactory device (located outside the vehicle, right), we used noise-canceling headphones for the sound output of the driving simulation.**

Important groundwork in the domain of AD is the study conducted by Beller, Heesen and Vollrath [6], who demonstrated the potential of a binary reliability display for AVs in a dual-task experiment. Since then, various papers have addressed reliability/uncertainty displays in the driving domain. Helldin et al. could show that such displays can also improve performance and comfort in Take-Over scenarios [25]. Recent studies have addressed potential metrics and design approaches for in-vehicle displays [49], but also augmented reality [40], or less obtrusive modalities such as haptics [39]. The presentation of different levels of reliability/uncertainty became more and more fine grained in these experiments, aiming to provide drivers more detailed information about the system state. However, a problem reliability displays share with any warning information is that, if (due to an offensive warning strategy) users face too many false alarms, they might simply ignore them ("cry wolf effect" [10]). Considering vehicle safety, drivers already seem to often ignore warning lights in vehicles [32].

As in the near future more and more potentially safety critical systems will be operated by everyday consumers [71], overtrust/overreliance is widely debated in the field of robotics [67] and AD [66]. For example, in a recent series of simulator studies conducted by Volvo, nearly 30% of drivers crashed in a provoked accident scenario, despite hands on the wheel and eyes on the road, and the authors conclude that more research is necessary to find out how system limitations can be communicated to drivers more effectively [66]. We claim that olfactory notifications could benefit the driver in such situations, as smell is a sense with a strong emotional component [2, 20, 27].

For example, Baron and Kalsher [5] proved that the scent of lemon increases alertness and the mood of the driver. The emotion-eliciting effect of scents is particularly useful in inducing mood changes because they are almost always experienced clearly as either pleasant or unpleasant [20]. For instance, Alaoui-Ismaïli et al. [1] used scents of vanillin and menthol to trigger positive emotions in their subjects (mainly happiness and surprise), as well as methyl methacrylate and propionic acid to trigger negative emotions (mainly disgust and anger). The scents of lemon, peppermint, rose, and lavender have been shown as efficient in improving the hedonic experiences of the user [16, 17, 55], whereas lemon and lavender have also demonstrated to be a good medium of conveying useful information in the context of driving and beyond [9, 16, 18, 19, 44]. On this aspect, it is essential to keep olfactory stimuli synchronized with other modalities [48]. Further, scents have already been proven to have a positive impact on driving performance/behavior. Martin and Cooper [45] showed that the scent of lemon can improve drivers' braking performance, while Dmitrenko et al. [16] demonstrated that the scents of lemon, peppermint, and lavender could help to reduce the number of errors. Further, scents of peppermint, rosemary, eucalyptus and lemon have been proven to be useful for keeping drowsy drivers awake [23, 28, 50, 73]. Scents could also help to remind drivers on certain driving-relevant activities, as the sense of smell is known to have a strong link with memories [9, 26, 58].

## 3 DRIVING SIMULATOR STUDY

To find out if olfactory displays can ease monitoring for drivers and thus provide a valuable extension of visual reliability displays, we conducted a dual-task study in a driving simulator. Participants had to drive in a semi-automated (level 2) vehicle while performing a detection-response task (DRT) on a smartphone (see Figure 1). To counter potential criticism of our experimental setting (smartphone usage or engagement in secondary tasks is strictly forbidden at level 2 driving in most countries), we want to emphasize that (1) many drivers engage in side activities (for example on mobile devices) despite given legislation [72], and (2) if successful, the underlying concept can be easily adapted to other levels of automation (for example to improve Take-Over requests [25]) or even different safety-critical systems.

### Method

Although recent studies on reliability displays often presented multiple levels [40, 49], we chose to utilize a binary display because of two reasons. First, we believed that for an initial evaluation, drivers should not need to distinguish between multiple levels of uncertainty, and second, we wanted to shift the principle from reliability to "responsibility". Currently, drivers utilizing ADSs remain the responsible control authority any time. However, to make AD successful in the future, vehicle manufacturers must start taking over responsibility for their vehicles' actions when driving in automated mode (this is a precondition to achieve driving automation above level 2 [12]). Thus, the binary display utilized in our study indicates either, that the vehicle itself takes over full responsibility for the dynamic driving task (green color, see Figure 2), or that the driver him/herself is responsible in case system reliability drops, indicating that a manual intervention is likely (red color, see Figure 2).

While performing the DRT, drivers had to monitor and intervene (if necessary) in the longitudinal control system of the AV. Participants could thereby rely on the given reliability information – as long as reliability was high (green), no manual intervention is necessary, in case of low reliability (red) a system malfunction is likely to occur. Drivers thus had to succeed in two tasks: (1) performing the DRT while (2) monitoring/intervening in longitudinal control of the vehicle. In randomized order, participants thereby faced the following conditions:

- ***Baseline* Condition**: No information about the longitudinal system's performance is presented, participants had to manually adjust their monitoring behavior.
- ***Reliability* Display**: The reliability for longitudinal control is presented to the user in form of a classic binary reliability display.
- ***Olfactory*-supported Reliability Display**: In addition to the visual information, an olfactory notification will be issued in case the reliability display changes its status (high to low and vice versa). Drivers thus can keep their visual attention on the DRT until they perceive the olfactory stimulus.


**Figure 2: Visual design of the utilized binary reliability display (left: high system reliability, right: low system reliability).**

### Measurements and Research Questions

For statistical investigation we conducted the following measurements: performance in the DRT (true and false positive rate), braking behavior/overrides of the longitudinal control system (number of manual interventions, average brake duration, and intensity), as well as subjective scales addressing user acceptance and trust. Therefore, we utilized the trust scale (TS) by Jian et al. [35] (which is widely used among trust researchers and provides sub-scales for both trust and distrust), and the Technology Acceptance Model (TAM) proposed by [13] (that assesses a user's intention to actually use a given system, determined by his/her perceived ease of use, perceived usefulness, and attitude towards using [it] [13]). Since product usage leads to positive/negative emotions [14], and odors have a strong emotional component [2], we also wanted to find out if the presented interface affects participants' emotional response. Therefore, we utilized the Positive/Negative Affect Scale (PANAS, [68]). Additionally, we conducted semi-structured interviews (assessing their perception of the system, potentially changed behavior, etc.) with all participants after the experiment. By statistical evaluation, we wanted to answer the following research questions:

- **RQ1:** Can olfactory notifications increase performance in the driving task (quantified by objective measurements assessing braking behavior)?
- **RQ2:** Can olfactory notifications increase performance in the side activity (detection-response task, quantified by true/false positive rates)?
- **RQ3:**How are olfactory notifications trusted and accepted by potential users in comparison with visual or the total absence of reliability displays (quantified by standardized scales such as TAM, TS, and PANAS)?

### Driving Task

We implemented our driving scenario (using IPG CarMaker) on the basis of Beller et a. [6], where drivers had to monitor an adaptive cruise control system. In our setting, the vehicle was driving on the left lane of a straight 2-lane highway segment with 120km/h. Every 30 seconds, the AV encountered a lead vehicle with a lower speed of just 70km/h, thus the system had to slow down, what was followed by the lead vehicle changing to the right lane (as soon as the ego-vehicle reached the same speed as the lead vehicle). Then, the ego-vehicle could accelerate again and continue driving with 120km/h for roughly 30 seconds, before the next lead vehicle appeared. We alternated phases of ca. 2 minutes (i.e., 4 lead vehicles) in either *high* (all vehicles detected and the AV slows down by itself) or *low* reliability (2 out of 4 cars not detected, where the driver had to intervene and brake manually). Each drive included 24 lead vehicles (roughly 12 minutes duration depending on participants' braking behavior during the three phases that require manual interventions) and thus 3 alternating phases of high and low reliability (in randomized order).

### Reliability Display and Olfactory Device

To communicate the reliability levels (*low* and *high*) we displayed either a green or red status symbol (see Figure 2) prominent on a tablet in the vehicle's center console (Google Pixel C, see Figure 1). The symbol changed after clearing the 4th vehicle of every section in the driving task to the new reliability level (condition *reliability*). In condition *olfactory*, we additionally communicated a change in reliability levels using two odors (lemon for a change to *low* and lavender for a change to *high* reliability). We used these two scents as both of them have been used to convey driving-relevant information in the past [16, 17]. Lemon was chosen for the change to *low* reliability, because it is known to keep the driver alert [5, 56] and to have an arousing effect on users [34]. Lavender was chosen for the switch to the *high* reliability level, because it is known to help drivers become aware of information they could have missed when relying only on visual stimuli [16], and as it is one of the most commonly used relaxing stimuli in olfactory research [3, 43, 46].
We presented these scents in an automated way, adapting a custom-made and fully controllable scent-delivery device (see [15] for design details). The device delivered the scented air from an air compressor (*Revell Masterclass*) attached to an air filter (5 micron filter from *Shako Co Ltd.*). The clean air was propelled through glass jars (using plastic tubes of 4mm in diameter) containing 6g of 100% pure essential oils (*"miaroma"* essential oils from *Holland & Barrett Int. Ltd.*) of lavender (*Lavandula officinalis*) and lemon (*Citrus limon*) with an air pressure of 1 bar. The scent-delivery nozzle (output) was located above the glove compartment, pointing towards the participant's face, approximately 1.5m away from the driver's nose (this distance could be shortened, if required by the application scenario [17]). The flow of air was controlled using electric valves (*SMC Compact Direct Operated 2 Port Solenoid Valves*) and an Arduino board [15]. The scent delivery was working with the vehicle's AC system being constantly on. All involved technical components

(smartphone for the DRT, scent delivery device, driving simulation, etc.) were synchronized by in-house software to enable time-critical measurements and repeatability (as suggested by [59]).

### Detection-Response Task

For the secondary task we implemented an HTML5/JavaScript application running on a OnePlus One 5.5" smartphone. On white background, each cell of a 3x3 grid was updated every second randomly showing numbers between 0 and 9 (or no number respectively, see Figure 1). Every time the number "6" appeared, participants had to press a large button at the bottom of the screen (once, a second button press was dismissed in this case). To evaluate performance in the detection-response task, we calculated the average response time for true positives, the true positive rate, and the false positive rate.

### Procedure

Prior to taking a part in the study, all participants were screened for potential olfactory dysfunctions or adverse reactions to string scents. Upon arrival, participants were given a consent form and the experimenter explained the experiment verbally to participants before starting the driving phase(s). Participants were encouraged to ask questions, if anything remained unclear. After a short test drive helping participants to get used to the driving simulator, the experiment started with one of the three conditions (*baseline* – no visual and no olfactory stimuli, *reliability* – with visual notifications involved, or *olfactory* – visual notifications combined with the olfactory stimuli).

The order of the conditions was quasi-randomized. Each condition lasted about 12 minutes and the switch between the reliability levels took place every two minutes. In the olfactory condition, the scent was triggered simultaneously with the switch between the visual stimuli and was delivered for five seconds. Participants were asked to complete a short questionnaire assessing demographics before the driving phases and the set of standardized scales after each condition. At the end of the experiment, we further conducted a five minutes long semi-structured interview with each participant.

## 4 RESULTS

In total, 25 participants aged between 19 and 38 years ($M = 24, SD = 3.98$ years, 10 female, 15 male) voluntarily participated in the study. Participants have reported to have no olfactory dysfunctions, adverse reactions to strong scents, respiratory problems or flu, and female participants confirmed that they are not pregnant. Participants were recruited on an opportunity-sampling basis and all expressed written consent. In the following, we present the results of our statistical evaluation. Effects are reported as statistically significant if $p < .05$, we used *IBM SPSS* Version 24 and (one way) repeated measures ANOVA (Greenhouse-Geisser in case Mauchly's test for sphericity failed) with Bonferroni correction and respectively Friedman ANOVA, if the data did not follow a normal distribution. A summary of descriptive statistics and evaluation results is presented in Tables 1 and 2.

### Objective Measures

*Driving Performance.* Considering driving performance, we calculated the number of brake pedal actuations, the average duration, as well as the average intensity of braking actions. All parameters showed no significant differences between the conditions (see Table 1 for descriptive statistics). Friedman ANOVA (test for normal distribution failed) resulted in $\chi^2(2) = .427, p = .080$ for the average number of brake actuations, and in $\chi^2(2) = .250, p = .882$ for the average duration of a braking action. A repeated measures ANOVA (assumptions for normal distribution and data sphericity met) for the average intensity of each braking action did not show significant differences ($F(2, 46) = 2.844, p = .068$) as well. However, pairwise comparisons using Bonferroni correction would have shown a difference between the conditions *baseline* and *olfactory* ($p = 0.02$, we report this fact as ANOVA just slightly missed the significance level of .05).

*Secondary Task Performance.* To assess the performance in the detection-response task, we evaluated the average reaction time for true positives, as well as the true and false positive rate (see Table 1 for descriptive statistics). Only the reaction time was not normally distributed, and there were no significant differences applying Friedman ANOVA ($\chi^2(2) = .75, p = .687$). For the true positive rate (TP), we found a significant difference using repeated measures ANOVA ($F(2, 46) = 3.496, p = .039$), however, post-hoc tests using Bonferroni correction showed no individual differences (if any, conditions *baseline* and *olfactory* were slightly above the significance level with $p = .63$, where *olfactory* showed the highest true positive rate). The false positive rate, on the other hand, resulted in a significant difference ($F(2, 46) = 4.823, p = .013$), where post-hoc tests revealed the origin between conditions *olfactory* and *baseline* ($p = .024$, where *olfactory* resulted in the lowest false positive rate).

### Subjective Measures

*Standardized Scales.* Reliability analysis showed acceptable values for Cronbach's alpha (above .722 or higher) for all sub-scales, thus we were able to calculate mean scale values. Considering the trust scale from Jian et al. [35], we found significant differences for both sub-scales of trust and distrust (average of the respective scale items). Distrust significantly differs with respect to the conditions (Greenhouse-Geisser

| | Condition Mean (SD) | | | Statistics | |
|---|---|---|---|---|---|
| | *Baseline* | *Reliability* | *Olfactory* | F | Sig ($\eta_p^2$) |
| Driving Behavior | | | | | |
| Nr. of brakes | 15.08 (12.89) | 13.67 (5.92) | 20.33 (26.89) | - | .808 (-) |
| Avg. brake duration (s) | 2.91 (2.48) | 2.46 (.92) | 2.33 (1.01) | - | .882 (-) |
| Avg. brake intensity | .52 (.19) | .49 (.18) | .46 (.20) | 2.844 | .068 (.11) |
| Secondary Task Performance | | | | | |
| Avg. response time TP (s) | .81 (.10) | .80 (.10) | .80 (.12) | - | .68 (-) |
| True positive rate | .62 (.09) | .63 (.09) | .65 (.09) | 3.496 | **.039 (.13)** |
| False positive rate | .0143 (.004) | .0125 (.005) | .0113 (.003) | 4.823 | **.013 (.17)** |

**Table 1: Descriptive and test statistics of objective data (braking behavior, secondary task performance). In case of missing F-values, Friedman ANOVA was utilized (test statistics reported throughout the results section). Significant differences are printed in bold face.**

since failed precondition for sphericity, $F(1.606, 38.550) = 18.508, p < .001$). Post-hoc analysis using Bonferroni correction showed differences between the conditions *baseline* and *olfactory* ($p < .001$), as well as *reliability* and *olfactory* ($p = .001$), but not between *baseline* and *reliability*. Contrarily, in the sub-scale trust ($F(1.454, 34.891) = 22.725, p < .001$), both conditions *olfactory* ($p < .001$) and *reliability* ($p = .001$) significantly differed from the *baseline*. However, no difference between *reliability* and *olfactory* was present here.

In the Technology Acceptance Model (TAM), we were able to find significant differences in all sub-scales. Perceived ease of use (PEOU) significantly differed in the result of the repeated measures ANOVA (Greenhouse-Geisser: $F(1.428, 34.273) = 37.061, p < .001$). Pairwise comparisons (Bonferroni) revealed that the *baseline* differs to both conditions *olfactory* ($p < .001$) and *reliability* ($p < .001$), while there were no differences between the latter two. Exactly the same result was obtained for perceived usefulness (PU, $F(1.348, 32.349) = 17.272, p < .001$). Also here, only the *baseline* differed to *olfactory* ($p = .001$) and *reliablity* ($p = .001$). Regarding the attitude towards using the system (ATT), all conditions have demonstrated significant differences, $F(1.484, 35.615) = 37.061, p < .001$. Condition *olfactory* was rated as highest and showed differences to *reliability* ($p = .01$) and *baseline* ($p < .001$), but also *reliability* was significantly higher than the *baseline* condition ($p = .018$). Since intention to use the system (INT) does not represent a scale variable, we utilized a non-parametric test (Friedman ANOVA, $\chi^2(2) = 19.3, p < .001$). Here, pairwise comparisons showed that only conditions *baseline* and *olfactory* significantly differed ($p = .002$) from each other. When looking at the results of the Positive/Negative Affect Scale (PANAS), we could not find any differences regarding positive affect (PA, preconditions for repeated measures ANOVA, as well as data sphericity met, $F(2, 48) = .280, p = .869$). The negative affect (NA) on the other hand resulted in significant differences ($F(2, 48) = .6.729, p = .003$), where post-hoc tests using Bonferroni correction revealed that condition *olfactory* had significantly less negative affect than *reliability* ($p = .036$) and *baseline* ($p = .02$), while there was no differences between *baseline* and *reliability*.

*Semi-Structured Interviews.* In the interviews, all 25 participants have confirmed that they have perceived the scents used in the experiment. They mainly emphasized the fact that the scents were intense enough to get perceived quickly and that this was helpful. For example, *P19* said: *"The scents were always quite intense in the beginning. It was good, because I could always understand when the switch between the reliability levels took place."* Also, all the 25 participant had experienced neither scent lingering, nor cross-contamination during the driving phase. They particularly liked that the timing of the scent-delivery was spot-on, that the scent disappeared quickly and matched the visual notifications very well. For example, *P12* said: *"The scents were so succinct that they appeared at the right time and were then gone relatively quickly."*

Scents were also perceived as helpful in performing the task of driving and in monitoring the autonomous system. 20/25 participants had mentioned the scents as helpful in perceiving the change between the reliability levels and as supportive in capturing the visual information displayed in the center console. For example, *P14* said: *"The scents helped, especially when there was no eye contact with the display."* Moreover, 19/25 participants admitted that they had to monitor the display less, thanks to the scents. They argued that they had to look on the display less, could rely on scents, and that their attention was grasped by the scents. For example, *P13* said: *"Thanks to the scents, when I was interacting with the phone, I was sure that with this system I can do anything."*

In terms of usefulness, 18/25 participants also mentioned that they find olfactory interaction generally useful in automotive context, considering that the choice of scents is

| | Condition Mean (SD) | | | Statistics | |
|---|---|---|---|---|---|
| | *Baseline* | *Reliability* | *Olfactory* | F | Sig ($\eta_p^2$) |
| Trust Scale | | | | | |
| Trust | 1.71 (1.08) | 2.99 (1.06) | 3.35 (1.26) | 22.725 | **<.001 (.486)** |
| Distrust | 3.93 (1.40) | 3.15 (1.08) | 2.28 (0.91) | 18.508 | **<.001 (.435)** |
| Technology Acceptance Model | | | | | |
| Perceived ease of use | 2.38 (1.28) | 4.06 (1.02) | 4.21 (0.92) | 37.061 | **<.001 (.607)** |
| Perceived usefulness | 1.32 (1.22) | 2.60 (1.41) | 2.91 (1.65) | 17.272 | **<.001 (.418)** |
| Attitude towards using the system | 1.92 (1.64) | 3.11 (1.42) | 3.84 (1.62) | 15.232 | **<.001 (.388)** |
| Intention to use the system | 1.44 (1.66) | 2.52 (1.61) | 3.20 (2.00) | - | **<.001 (-)** |
| Positive and Negative Affect Scale | | | | | |
| Positive Affect | 2.82 (1.10) | 2.74 (1.08) | 3.07 (1.20) | 2.038 | .141 (.078) |
| Negative Affect | 2.49 (1.57) | 2.24 (1.34) | 1.67 (1.34) | 6.729 | **.003 (.219)** |

**Table 2: Descriptive and test statistics of subjective scales (Trust scales, Technology Acceptance Model, Positive and Negative Affect Scale). In case of missing F-values, Friedman ANOVA was utilized (test statistics reported throughout the results section). Significant differences are printed in bold face.**

performed carefully, appropriate training is carried out, and the scent-delivery is well controlled. For example, *P11* said: *"It's a good idea, but you need to be careful about the choice of scents."*, whereas *P13* said: *"It's something brand new! It was very nice! You just need to be careful that not too many scents are used... 2-3 very different scents would be good, I think."*
At the end of the interview, we encouraged participants to suggest further scenarios, in which they consider olfactory feedback to be effective. 5/25 participants recommended using scents as warnings for such non-urgent notifications as a traffic jam or a bad weather alert, and a low petrol level notification. For example, *P21* said: *"I would use scents when there is enough time, when I can decide what I can take over."* Another 5/25 participants suggested rather using scents for safety critical notifications (also as a support to visual stimuli), such as ACC, inter-vehicle distance, and vehicles passing by on the left/right. For instance, *P4* said: *"Scents could come when you drive too close to a car in front of you, when a traffic light goes red, or when a child crosses the road."* Furthermore, 3/25 participants expressed a wish for scents to convey vehicle diagnosis-related data. *P11* mentioned *"engine overheating"*, *P15* an *"oil leak"*, and *P19* generalized this to *"problems with the car"*. Two participants decided that scents are good for *"take a break"* notifications, e.g. *P7* said: *"It makes sense to use scents in the car, because they make the driver awake on short-term."* Finally, 4/25 participants referred to the well-being of the driver, e.g. *P21* said: *"When I get into the car and it smells nice, it contributes to comfort, of course."*

### Summary of Results

Only the combination of visual and olfactory reliability information (condition *olfactory*) showed differences in driving behavior (less intensive brake pedal actuations) and secondary task performance (lower false positive rate), while the provision of the visual display only (condition *reliability*) did not result in an improvement compared to the *baseline* drive. In subjective scales (TS, TAM, PANAS) a significant difference was visible for both test conditions (*reliability* and *olfactory*) compared to the baseline in most sub-scales, while *olfactory* notifications showed significantly less distrust (TS) and negative affect (PANAS), as well as a higher attitude towards using the system (TAM) compared to visual *reliability* information only. Participants' positive attitude towards *olfactory* notifications was further confirmed in semi-structured interviews.

## 5 DISCUSSION

The results of our study provide multiple interesting insights and confirm, in general, the potential of olfactory notifications for trust calibration. Regarding driving behavior (**RQ1**), a statistically significant difference was only present between the condition *olfactory* and the *baseline* (participants showed less average braking intensity and thus braked "smoother"). This difference is only visible in post-hoc tests, while the overall ANOVA result slightly missed meeting the significance level. Although it is often emphasized that pairwise comparisons should only be conducted in case of significant ANOVA results, there is also the view that pairwise differences can be seen as valid even if the global effect is not significant [33]. However, we believe it is necessary to include a larger sample size to either confirm or reject the observed tendency. At the present stage, our study cannot fully confirm the results of Beller et al. [6] regarding strong significant differences in braking behavior.
Considering **RQ2**, the combination of visual and olfactory reliability display significantly decreased the false positive

rate in the secondary task (visual detection-response task) compared to the baseline condition. There also exist some tendencies that condition *olfactory* resulted in a higher true positive rate as compared to the *baseline* condition (ANOVA result significant but post-hoc not, however Bonferroni is known to be conservative in pairwise comparisons [70]). The provision of a visual reliability display only did not significantly improve secondary task performance compared to the baseline, what highlights the potential benefit of olfactory notifications presented along the visual stimuli.

Regarding standardized subjective scales assessing trust and user acceptance (**RQ3**), we can report increased trust and acceptance towards olfactory notifications. Condition *olfactory* induced significantly less distrust and received significantly higher attitude towards using the system, compared to visual-only provision of reliability. Also, the provision of *olfactory* cues resulted in a significantly lower negative affect in PANAS, thus this modality was not perceived negatively among study participants. Subjects' positive attitude towards olfactory notifications was further confirmed in semi-structured interviews.

Our observations are in line with the previous findings on multimodal in-car interfaces, where drivers were shown to perform better when assisted by notifications consisting of multiple modalities (e.g., as per [16, 53]). Study results confirm this in the scope of AD and trust in automation. Our findings are also matching the evidence found in the fields of psychology and neuroscience, where the sense of smell has been demonstrated as an efficient medium of conveying semantically congruent cues [24]. Some might question that olfactory stimuli meet the requirements of time-sensitive notifications, but while vision, audition, and touch traverse the perceived information en route from periphery to primary cortex, in olfaction information reaches the primary cortex directly [54], what is a unique advantage compared to other senses. Further, it is known that a high odorant concentration can spontaneously shift our attention to olfaction [61].

Still, we do not suggest combining olfactory notifications with other modalities permanently, and we would neither suggest the same for any additional form of reliability display. We rather emphasize that olfactory notifications (as other modes of communication such as haptics [39]) are a valuable extension to be included in intelligent user interfaces: Olfactory stimuli could be applied as feedback messages, in cases when visual notifications are likely to be missed [16]. As it cannot be guaranteed that drivers can be reached with any given modality due to their engagement in arbitrary secondary activities (that will vary with respect to the demand of different perceptional channels), future interfaces should become context-aware and thereby take both environmental/operational properties of the situation, as well as personal preferences [51] of different drivers into account.

The National Transportation Safety Board (NTSB) reported that in the fatal Tesla accident in 2016, the driver did not respond to multiple visual and auditory warnings issued by the system prior to the accident. We do not claim that olfactory notifications would have made the difference in this situation, but we want to raise the question: what could have happened, if a strong scent, for example the smell of a broken engine, would have been issued to gain the driver's attention [57]? An answer to this question, as well as when and how the support of olfactory notifications yields the best results, should be addressed in future studies and detailed research. However, our study provides initial insights that highlight the potential of olfactory notifications for trust calibration in AVs.

## 6 LIMITATIONS AND FUTURE WORK

As our findings demonstrate promising tendencies in terms of olfactory enhanced reliability displays, it is worth exploring multiple levels of reliability conveyed by scents in the future. This could be achieved by either using two scents of different intensity levels (e.g., as in [15]) or by extending the range of scents (e.g., as in [16, 17]) and assigning a certain scent to every urgency level (e.g., as in [44]). When working with scents, it is important to acknowledge the subjective element of scent perception. For example, four of our participants said in the post-experiment interview that they did not like the scent of lavender. In the future, it would be necessary to explore customizable olfactory interfaces, allowing participants to select the scent of their preference. Also, as the selection of the scents might not work for everyone and in every situation (e.g., not in case of a flu), it would be a good idea to explore other modalities, such tactile [8, 62] and ambient light [42] interfaces for conveying automation reliability-relevant information. We have tested our olfactory interface in a high-fidelity driving simulator, with a real car interior. Interviews conducted with the participants have revealed no scent lingering or cross-contamination artifacts experienced during the experiments. This suggests that the interface is suitable for the use in a real car. However, this would need to be supported by further studies in the real road environment and for an extended time frame. The location of the scent-delivery nozzle and the interference of the scented airflow with the vehicle's AC (or air coming through an open window) would need to be investigated further. This might include positioning the nozzle closer to the driver's nose (as per [15]) or temporarily replacing olfactory stimulation by other modalities (e.g., touch [62]). On-road studies would also help understand how do drivers feel about using scents over a longer period of time, how their sensitivity to the olfactory stimuli changes over time, and if scents get absorbed by the car's interior on long term. Furthermore, such explorations could reveal the efficiency of the olfactory stimuli in the presence of external scents (e.g., coffee or a dog on the rear seat). In terms of neutralizing the delivered scents, it would be useful to investigate different ventilation parameters (as per [15]) and to explore the application of the "olfactory white" [69].

## 7 CONCLUSION

In this paper, we have evaluated the potential of an added olfactory UI in supporting reliability displays for trust calibration in automated vehicles. Results of a driving simulator study (N=25) comparing three conditions (visual reliability display only, visual display supported by olfactory notifications, baseline condition without any reliability information) confirm our assumption that olfactory cues can improve performance in a dual-task setting. We can report tendencies that, with support of olfactory cues, participants showed smoother braking behavior compared to the baseline condition (quantified as brake pedal actuations during manual interventions in case the longitudinal control system failed). Also, adding olfactory notifications to the visual stimuli resulted in significantly higher performance in the visual detection-response task. In addition, this (at least in the context of trust calibration) yet unused modality was subjectively preferred by study participants based on subjective evaluation. Participants rated the system with added olfactory cues significantly better in sub-scales of the Technology Acceptance Model (TAM) [64], the Trust Scale [35], and the Positive/Negative Affect Scale (PANAS) [68]. Participants' positive attitude towards olfactory notifications was further confirmed in semi-structured interviews, where 80% of the participants stated that olfactory cues are helpful in perceiving a change in vehicle reliability levels. Overtrust is an issue that already led to (even fatal) accidents with automated vehicles [67, 71] and could hinder a success of the automated driving technology. Identifiying additional methods to calibrate trust is, thus, timely and important [66]. Olfactory cues could become a valuable asset helping to regain attention of drivers that are engaged in secondary tasks (and thus out of the loop), allowing them to more reliably assess and react to unknown circumstances.

## ACKNOWLEDGMENTS

We applied the FLAE approach for the sequence of authors. This work is supported under the FH-Impuls program of the German Federal Ministry of Education and Research (BMBF) under Grant Number 13FH7I01IA (SAFIR) and by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program under the grant agreements No 638605 and No 737576.

## REFERENCES

[1] O. Alaoui-Ismaïli, O. Robin, H. Rada, A. Dittmar, and E. Vernet-Maury. 1997. Basic emotions evoked by odorants: comparison between autonomic responses and self-evaluation. *Phys. & Beh.* (1997).

[2] J. Amores and P. Maes. 2017. Essence: Olfactory Interfaces for Unconscious Influence of Mood and Cognitive Performance. In *CHI '17*. ACM, New York, NY, USA.

[3] T. Atsumi and K. Tonosaki. 2007. Smelling lavender and rosemary increases free radical scavenging activity and decreases cortisol level in saliva. 150 (03 2007).

[4] Lisanne Bainbridge. 1983. Ironies of automation. In *Analysis, Design and Evaluation of Man–Machine Systems 1982*. Elsevier, 129–135.

[5] R. A. Baron and M. J. Kalsher. 1998. Effects of a Pleasant Ambient Fragrance on Simulated Driving Performance: The Sweet Smell of... Safety? *Env. and Beh.* (1998).

[6] Johannes Beller, Matthias Heesen, and Mark Vollrath. 2013. Improving the driver–automation interaction: An approach using automation uncertainty. *Human factors* 55, 6 (2013), 1130–1141.

[7] Adam Bodnar, Richard Corbett, and Dmitry Nekrasovski. 2004. AROMA: Ambient Awareness Through Olfaction in a Messaging Application. In *Proceedings of the 6th International Conference on Multimodal Interfaces (ICMI '04)*. ACM, New York, NY, USA, 183–190. https://doi.org/10.1145/1027933.1027965

[8] S.S. Borojeni, T. Wallbaum, W. Heuten, and S. Boll. 2017. Comparing Shape-Changing and Vibro-Tactile Steering Wheels for Take-Over Requests in Highly Automated Driving. In *AutomotiveUI '17*. ACM.

[9] S. Brewster, D. McGookin, and C. Miller. 2006. Olfoto: Designing a Smell-based Interaction. In *CHI '06*. ACM, New York, NY, USA.

[10] Shlomo Breznitz. 2013. *Cry wolf: The psychology of false alarms*. Psychology Press.

[11] U. Castiello, G.M. Zucco, V. Parma, C. Ansuini, and R. Tirindelli. 2006. Cross-Modal Interactions between Olfaction and Vision When Grasping. *Chemical Senses* 31, 7 (2006).

[12] SAE On-Road Automated Vehicle Standards Committee. 2014. Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems.

[13] Fred D Davis. 1993. User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. *International journal of man-machine studies* 38, 3 (1993), 475–487.

[14] Pieter MA Desmet and Paul Hekkert. 2007. Framework of product experience. *International journal of design* 1, 1 (2007), 57–66.

[15] D. Dmitrenko, E. Maggioni, and M. Obrist. 2017. OSpace: Towards a Systematic Exploration of Olfactory Interaction Spaces. In *ISS '17*. ACM, New York, NY, USA.

[16] D. Dmitrenko, E. Maggioni, and M. Obrist. 2018. I Smell Trouble: Using Multiple Scents To Convey Driving-Relevant Information. In *ICMI '18*. ACM, New York, NY, USA.

[17] D. Dmitrenko, E. Maggioni, C.T. Vi, and M. Obrist. 2017. What Did I Sniff?: Mapping Scents Onto Driving-Related Messages. In *AutomotiveUI '17*. ACM, New York, NY, USA.

[18] D. Dmitrenko, C.T. Vi, and M. Obrist. 2016. A Comparison of Scent-Delivery Devices and Their Meaningful Use for In-Car Olfactory Interaction. In *AutomotiveUI '16*. ACM, New York, NY, USA.

[19] D. Dobbelstein, S. Herrdum, and E. Rukzio. 2017. inScent: A Wearable Olfactory Display As an Amplification for Mobile Notifications. In *ISWC '17*. ACM, New York, NY, USA.

[20] H. Ehrlichman and L. Bastone. 1992. Olfaction and emotion. In *Science of olfaction*. Springer.

[21] Mica R Endsley. 2017. Autonomous driving systems: a preliminary naturalistic study of the Tesla Model S. *Journal of Cognitive Engineering and Decision Making* 11, 3 (2017), 225–238.

[22] Richard Finger and Ann M Bisantz. 2002. Utilizing graphical formats to convey uncertainty in a decision-making task. *Theoretical Issues in*

*Ergonomics Science* 3, 1 (2002), 1–25.

[23] H. Funato, M. Yoshikawa, M. Kawasumi, S. Yamamoto, M. Yamada, and Y. Yanagida. 2009. Stimulation effects provided to drivers by fragrance presentation considering olfactory adaptation. In *Intelligent Vehicles Symposium 2009*. IEEE.

[24] J.A. Gottfried and R.J. Dolan. 2003. The Nose Smells What the Eye Sees: Crossmodal Visual Facilitation of Human Olfactory Perception. *Neuron* 39, 2 (2003).

[25] Tove Helldin, Göran Falkman, Maria Riveiro, and Staffan Davidsson. 2013. Presenting system uncertainty in automotive UIs for supporting trust calibration in autonomous driving. In *Proceedings of the 5th international conference on automotive user interfaces and interactive vehicular applications*. ACM, 210–217.

[26] R.S. Herz and T. Engen. 1996. Odor memory: Review and analysis. *Psychonomic Bulletin & Review* 3, 3 (1996).

[27] R. S. Herz. 2002. Influences of Odors on Mood and Affective Cognition. In *Olfaction, Taste, and Cognition*. Cambridge University Press.

[28] S. Hiroike, S. Doi, T. Wada, E. Kobayashi, M. Karaki, N. Mori, T. Kusaka, and S. Ito. 2009. Study of olfactory effect on individual driver under driving. In *ICME '09*.

[29] Kevin Anthony Hoff and Masooda Bashir. 2015. Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors* 57, 3 (2015), 407–434.

[30] R.W. Holland, M. Hendriks, and H. Aarts. 2005. Smells Like Clean Spirit Nonconscious Effects of Scent on Cognition and Behavior. *Psychological Science* 16, 9 (2005).

[31] T. Hongratanaworakit. 2008. Relaxing Effect of Rose Oil on Humans. *Natural Product Communications* (2008).

[32] HPI. 2018. How long do you leave your dashboard lights on? Online. Retrieved October 01, 2018 from https://goo.gl/dMf99W.

[33] Jason Hsu. 1996. *Multiple comparisons: theory and methods*. Chapman and Hall/CRC.

[34] J. Ilmberger, E. Heuberger, C. Mahrhofer, H. Dessovic, D. Kowarik, and G. Buchbauer. 2001. The influence of essential oils on human attention. I: Alertness. *Chemical Senses* 26, 3 (2001).

[35] Jiun-Yin Jian, Ann M Bisantz, and Colin G Drury. 2000. Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics* 4, 1 (2000), 53–71.

[36] L.M. Kay. 2011. Olfactory Coding: Random Scents Make Sense. *Current Biology* 21, 22 (2011).

[37] J.K. Kiecolt-Glaser, J.E. Graham, W.B. Malarkey, K. Porter, S. Lemeshow, and R. Glaser. 2008. Olfactory influences on mood and autonomic, endocrine, and immune function. *Psychoneuroendocrinology* 33, 3 (2008).

[38] Jeamin Koo, Jungsuk Kwac, Wendy Ju, Martin Steinert, Larry Leifer, and Clifford Nass. 2015. Why did my car just do that? Explaining semi-autonomous driving actions to improve driver understanding, trust, and performance. *International Journal on Interactive Design and Manufacturing (IJIDeM)* 9, 4 (2015), 269–275.

[39] Alexander Kunze, Steve Summerskill, Russell Marshall, and Ashleigh J Filtness. 2018. Preliminary evaluation of variables for communicating uncertainties using a Haptic seat. (2018).

[40] Alexander Kunze, Stephen J Summerskill, Russell Marshall, and Ashleigh J Filtness. 2018. Augmented Reality Displays for Communicating Uncertainty Information in Automated Driving. In *Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM, 164–175.

[41] John D Lee and Katrina A See. 2004. Trust in automation: Designing for appropriate reliance. *Human factors* 46, 1 (2004), 50–80.

[42] A. Löcken, W. Heuten, and S. Boll. 2015. Supporting Lane Change Decisions with Ambient Light. In *AutomotiveUI '15*. ACM, NY, USA.

[43] H.W. Ludvigson and T.R. Rottman. 1989. Effects of ambient odors of lavender and cloves on cognition, memory, affect and mood. *Chem. Sen.* (1989).

[44] E. Maggioni, R. Cobden, D. Dmitrenko, and M. Obrist. 2018. Smell-O-Message: integration of olfactory notifications into a messaging application to improve users' performance. In *ICMI '18*. ACM, New York, NY, USA.

[45] G. N. Martin and J. A. Cooper. 2007. Adding zest to difficult journeys: Odour effects on simulated driving performance. In *BPS Annual Conference*.

[46] M. Moss, J. Cook, K. Wesnes, and P. Duckett. 2003. Aromas of rosemary and lavender essential oils differentially affect cognition and mood in healthy adults. *Int. J. of Neuroscience* (2003).

[47] Bonnie M Muir. 1987. Trust between humans and machines, and the design of decision aids. *International Journal of Man-Machine Studies* 27, 5-6 (1987), 527–539.

[48] N. Murray, G.M. Muntean, Y. Qiao, and B. Lee. 2018. *Olfaction-Enhanced Multimedia Synchronization*. Springer International Publishing.

[49] Brittany E. Noah, Thomas M. Gable, Shao-Yu Chen, Shruti Singh, and Bruce N. Walker. 2017. Development and Preliminary Evaluation of Reliability Displays for Automated Lane Keeping. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '17)*. ACM, New York, NY, USA, 202–208. https://doi.org/10.1145/3122986.3123007

[50] C. Oshima, A. Wada, H. Ando, N. Matsuo, S. Abe, and Y. Yanigada. 2007. Improved delivery of olfactory stimukus to keep drivers awake. In *Workshop on DSP for in-Vehicle and Mobile Systems*.

[51] Shimei Pan, Oliver Brdiczka, Giuseppe Carenini, Duen Horng Chau, and Per Ola Kristensson. 2016. What's Hot in Intelligent User Interfaces.. In *AAAI*. 4333–4334.

[52] Raja Parasuraman and Victor Riley. 1997. Humans and automation: Use, misuse, disuse, abuse. *Human factors* 39, 2 (1997), 230–253.

[53] I. Politis, S. Brewster, and F. Pollick. 2015. To Beep or Not to Beep?: Comparing Abstract Versus Language-Based Multimodal Driver Displays. In *CHI '15*. ACM, New York, NY, USA.

[54] J.L. Price. 1990. Olfactory system. *The human nervous system* (1990).

[55] N. Ranasinghe, P. Jain, N. Thi Ngoc Tram, K.C.R. Koh, D. Tolley, S. Karwita, L. Lien-Ya, Y. Liangkun, K. Shamaiah, C. Eason Wai Tung, C.C. Yen, and E.Y. Do. 2018. Season Traveller: Multisensory Narration for Enhancing the Virtual Reality Experience. In *CHI '18*. ACM, USA.

[56] B. Raudenbush, R. Grayhem, T. Sears, and I. Wilson. 2009. Effects of Peppermint and Cinnamon Odor Administration on Simulated Driving Alertness, Mood and Workload. *N.A.J. of Psychology* (2009).

[57] Andreas Riener. 2010. *Sensor-actuator supported implicit interaction in driver assistance systems*. Springer.

[58] F.R. Schab and R.G. Crowder. 2014. *Memory for odors*. Psychology Press.

[59] Clemens Schartmüller, Philipp Wintersberger, and Andreas Riener. 2017. Interactive Demo: Rapid, Live Data Supported Prototyping with U.S.E.. In *Adjunct Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM, Oldenburg, 239–241. https://doi.org/10.1145/3131726.3132046

[60] A. Seigneuric, K. Durand, T. Jiang, J.-Y. Baudouin, and B. Schaal. 2010. The nose tells it to the eyes: Crossmodal associations between olfaction and vision. *Perception* 39, 11 (2010).

[61] L. Sela and N. Sobel. 2010. Human olfaction: a constant state of change-blindness. *Experimental Brain Research* (2010).

[62] G. Shakeri, A. Ng, J.H. Williamson, and S.A. Brewster. 2016. Evaluation of Haptic Patterns on a Steering Wheel. In *Automotive'UI 16*. ACM, New York, NY, USA.

[63] Matthew Turk. 2014. Multimodal interaction: A review. *Pattern Recognition Letters* 36 (2014), 189–195.

[64] Viswanath Venkatesh, Michael G Morris, Gordon B Davis, and Fred D Davis. 2003. User acceptance of information technology: Toward a unified view. *MIS quarterly* (2003), 425–478.

[65] E. Vernet-Maury, O. Alaoui-Ismaïli, A. Dittmar, G. Delhomme, and J. Chanel. 1999. Basic emotions induced by odorants: A new approach based on autonomic pattern results. 75 (03 1999).

[66] Trent W Victor, Emma Tivesten, Pär Gustavsson, Joel Johansson, Fredrik Sangberg, and Mikael Ljung Aust. 2018. Automation Expectation Mismatch: Incorrect Prediction Despite Eyes on Threat and Hands on Wheel. *Human factors* (2018), 0018720818788164.

[67] Alan R Wagner, Jason Borenstein, and Ayanna Howard. 2018. Overtrust in the robotic age. *Commun. ACM* 61, 9 (2018), 22–24.

[68] David Watson, Lee Anna Clark, and Auke Tellegen. 1988. Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology* 54, 6 (1988), 1063.

[69] T. Weiss, K. Snitz, A. Yablonka, R.M. Khan, D. Gafsou, E. Schneidman, and N. Sobel. 2012. Perceptual convergence of multi-component mixtures in olfaction implies an olfactory white. *Proceedings of the National Academy of Sciences* (2012).

[70] Eric W Weisstein. 2004. Bonferroni correction. (2004).

[71] Philipp Wintersberger and Andreas Riener. 2016. Trust in technology as a safety aspect in highly automated driving. *i-com* 15, 3 (2016), 297–310. https://doi.org/doi:10.1515/icom-2016-0034

[72] Philipp Wintersberger, Andreas Riener, Clemens Schartmüller, Anna-Katharina Frison, and Klemens Weigl. 2018. Let Me Finish before I Take Over: Towards Attention Aware Device Integration in Highly Automated Vehicles. In *Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM, 53–65.

[73] M. Yoshida, C. Kato, Y. Kakamu, M. Kawasumi, H. Yamasaki, S. Yamamoto, T. Nakano, and M. Yamada. 2011. Study on Stimulation Effects for Driver Based on Fragrance Presentation. In *MVA2011 IAPR Conference on Machine Vision Applications*.